Education and debate

Numbers needed to treat derived from meta-analyses— sometimes informative, usually misleading

Liam Smeeth, Andy Haines, Shah Ebrahim

Department of Primary Care and Population Sciences, Royal Free and University College Medical School, University College London, London NW3 2PF
Liam Smeeth, *research fellow*
Andy Haines, *professor of primary care*

MRC Health Services Research Collaboration, Department of Social Medicine, University of Bristol, Bristol BS8 2PR
Shah Ebrahim, *professor of epidemiology and ageing*

Correspondence to: Professor Ebrahim shah.ebrahim@ bristol.ac.uk

The number needed to treat—the number of patients who must be treated to prevent one adverse outcome—is a widely used measure.[1] [2] It is increasingly being calculated by pooling absolute risk differences in trials included in meta-analyses.[3] [4] This option is available in statistical software and the *Cochrane Database of Systematic Reviews*.[5] [6] In this paper, we examine pooled numbers needed to treat derived from trials and meta-analyses of interventions to prevent cardiovascular disease. We show that a pooled number needed to treat may be misleading because of variation in the event rates in trials, differences in the outcomes considered, effects of secular trends on disease risk, and differences in clinical setting. The number needed to treat should be derived by applying the relative risk reductions from treatment which have been estimated by trials or meta-analysis to relevant baseline risks for different types of patients. This provides a range of possible numbers needed to treat in different patient groups.

Methods

The interventions selected for study were use of statins for lowering cholesterol concentrations in primary and secondary prevention of coronary heart disease.[7–11] In secondary prevention after myocardial infarction, the interventions considered were antiplatelet drugs,[12] β blockers,[13] and multiple risk factor interventions.[14] Comparisons of estimates of efficacy between clinical settings (that is, primary care compared with hospital clinics) were made using meta-analyses of trials of antihypertensive drugs in elderly people.[15] Secular trends from 1975 to 1995 in coronary heart disease mortality for men and women aged 55-64 years in England and Wales were obtained from routine mortality statistics.[16] Pooled absolute risk reductions were obtained using EasyMa software,[17] which also provides a pooled number needed to treat to avoid each event considered. Numbers needed to treat were calculated for five years of treatment.

Effect of choice of outcome

The effects of treatment with statins are shown in table 1. Numbers needed to treat vary greatly, depending on the outcome chosen. In communicating a positive message, it is tempting to chose the smallest number needed to treat, for example, "all the bad things that can happen."[18] The combined end point, all vascular events,

Summary points

Numbers needed to treat are often used to summarise treatment effects in a clinically relevant way

They are derived from the baseline risk without treatment and the reduction in risk achieved with treatment

Numbers needed to treat are sensitive to factors that change the baseline risk such as the outcome considered, patients' characteristics, secular trends in incidence and case fatality, and the clinical setting

Pooled numbers needed to treat derived from meta-analyses can be seriously misleading because the baseline risk often varies appreciably between the trials

Applying the pooled relative risk reductions calculated from meta-analyses or individual trials to the baseline risk relevant to specific patient group produces a useful number needed to treat

was made up of different proportions of events in the different trials. In the WOSCOPS and AFCAPS/ TexCAPS studies, coronary heart disease deaths made up 16% and 3% respectively of the combined end point.[7] [8]

In the AFCAPS/TexCAPS trial of lovastatin, the absolute coronary heart disease mortality difference was very close to zero, with 95% confidence intervals which included the possibility of benefit and also of harm. As the absolute difference comes close to zero, the number needed to treat becomes very large and approaches infinity. If the absolute difference is greater than zero, treatment is not beneficial, and the reciprocal of the absolute difference becomes a number needed to harm.[19]

Effect of variation in baseline risk

Interventions for secondary prevention after myocardial infarction are shown in table 2. While all treatments show very similar reductions in relative risk,

Table 1 Major trials of statins for coronary heart disease (CHD); rate ratios and numbers needed to treat

Trial	No of participants	Baseline risk of CHD mortality/100 person years	Rate ratios			Number needed to treat (5 years)		
			Total mortality	CHD mortality	All cardiovascular events	Total mortality	CHD mortality	All cardiovascular events
Primary prevention								
AFCAPS/TexCAPS[7]	6605	0.1	1.04	1.36	0.69	167*	1000*	28
WOSCOPS[8]	6595	0.4	0.78	0.67	0.70	118	182	28
Secondary prevention								
Scandinavian simvastatin survival study trial[9]	4444	1.6	0.71	0.59	0.64	33	31	8
CARE[10]	4159	1.2	0.92	0.81	0.75	133	95	11
Long-term intervention with pravastatin in ischaemic disease[11]	9014	1.4	0.78	0.77	0.80	41	64	17
Pooled effects (95% CI)			0.80 (0.74 to 0.87)	0.73 (0.66 to 0.81)	0.74 (0.71 to 0.77)	113 (77 to 285)	500 (222 to -)†	20 (17 to 25)

*AFCAPS/TexCAPS study reported a non-significant increased total and CHD mortality in the intervention group. Numbers needed to treat are derived from the lower limit of the 95% CIs of the risk differences in event rates to illustrate the lower limit within which the numbers might lie.
†No upper number needed to treat can be calculated as the upper 95% CI of pooled absolute risk difference is greater than zero. In these circumstances the number needed to treat is a number needed to harm.

the numbers needed to treat vary much more and have wide 95% confidence intervals. The baseline mortality in the individual trials varied greatly, in one case by more than an order of magnitude. These very large differences in absolute risk in the individual trials reflect the participants selected for inclusion. In the trials of multiple risk factor interventions, applying the same relative risk reduction to the trials with the highest and lowest coronary heart disease mortality rates resulted in numbers needed to treat for five years of 2 and 317 respectively.

Effect of secular trends

In primary prevention, the background population risk of coronary heart disease mortality is relevant in determining the numbers needed to treat. To illustrate the effect of secular trends on the number needed to treat, we have assumed that statins were available over the past 20 years and that the reductions in relative risk seen in the trials were achievable, applied to the whole population, and occurred over the whole range of coronary heart disease mortality (table 3). Numbers needed to treat for five years show instability over time. This is more noticeable for women than men, reflecting the greater absolute falls in mortality among women. If there are secular trends in case fatality, perhaps through changes in disease severity, caution in the use of numbers needed to treat may also apply to treatments used in secondary prevention.

Effect of clinical setting

For each outcome, the pooled relative risks are similar for primary care and secondary care trials. By contrast, the numbers needed to treat for five years vary twofold depending on the setting, but the 95% confidence intervals are wide. Despite the large numbers of patients in these trials, it is uncertain whether the large differences in numbers needed to treat are simply due to chance. Generalising the numbers needed to treat derived from trials undertaken in one setting to patient care in another setting may be misleading. However, the relative estimates of efficacy varied very little across the different settings and could be generalised with more confidence (table 4).

Discussion

Randomised controlled trials aim to achieve internal validity through careful inclusion criteria for participants, random allocation to intervention, and blind ascertainment of end points. Consequently, the absolute event rates in trials may bear little relation to the event rates that might be expected in routine clinical practice.

Pooling assumptions

There are two main categories of statistical model for meta analysis—fixed and random effect models. The fixed effect models assume that all the studies are estimating the same "true" effect, and that the variation in effects seen between different trials is a result of random error alone. The random effects model assumes that the trials included in the meta-analysis are a random sample from a hypothetical population of trials of different underlying effect sizes.[20 21] There will almost always be differences in the baseline risks of trials carried out in different populations and at different times. Consequently, there is no single, true, absolute risk difference, as assumed in the fixed effects models.

Table 2 Effects of different interventions as secondary prevention on mortality from coronary heart disease

Intervention	Pooled relative risk reduction	Range of mortality*in control groups/100 patient years	Pooled NNT for 5 years (95% CI)	Pooled absolute risk difference per 100 patient years (95% CI)
Antiplatelet drugs[12]	0.15	2.8 to 13.9	50 (25 to 500)	−0.4 (−0.8 to −0.04)
β blockers[13]	0.22	2.6 to 23.2	20 (13 to 33)	−1.0 (−1.5 to −0.6)
Statins[9-11]	0.26	0.5 to 2.6	100 (67 to 200)	−0.2 (−0.3 to −0.1)
Multiple risk factor intervention[14]	0.19	0.3 to 10.0	67 (29 to -)†	−0.3 (−0.7 to 0.1)

NNT=number needed to treat.
*From cardiovascular disease or coronary heart disease.
†No upper 95% CI calculated as absolute difference greater than zero.

Table 3 Secular trends in coronary heart disease (CHD) mortality in people aged 55-64 years and numbers needed to treat (NNT) with statins to avoid death from CHD

Year	CHD mortality/100 000		NNT for 5 years*	
	Men	Women	Men	Women
1975	7335	2116	83	286
1985	6700	2011	90	301
1995	3934	563	154	1075

*Assuming all patients are treated.

Table 4 Drug treatment of hypertension in elderly people in primary and secondary care[15]; rate ratios and numbers needed to treat (NNT)

Setting	Total mortality			Cardiovascular morbidity and mortality		
	No of participants	Rate ratio (95% CI)	NNT (95% CI)	No of participants	Rate ratio (95% CI)	NNT (95% CI)
Primary care	6 907	0.91 (0.80 to 1.03)	86 (35 to NA)	6 907	0.75 (0.67 to 0.84)	23 (16 to 38)
Secondary care	4 113	0.86 (0.79 to 0.94)	42 (27 to 109)	8 738	0.71 (0.66 to 0.78)	13 (11 to 18)
All trials	21 020	0.88 (0.82 to 0.94)	51 (34 to 118)	15 645	0.73 (0.68 to 0.78)	17 (14 to 21)

NA=not calculated as rate ratio includes possibility of harm.
Numbers differ because specific outcome data were not available for all trials.

Neither is the variation in risk difference between trials solely the result of a sampling effect. Decisions affecting the baseline risk of participants in a trial, such as inclusion and exclusion criteria or geographical setting, are not made in a random way, as is assumed in the random effects models. Consequently, pooled numbers needed to treat may contravene statistical assumptions made by these models.

Duration of treatment effect
Trials have different lengths of follow up. However, to produce a number needed to treat for five years, for example, all the absolute risk differences need to be standardised for five years. This standardisation requires an assumption of constancy of effect over time, an assumption that may not be reasonable. For example, in the Scandinavian simvastatin survival study, no effect on total mortality was evident until one year of treatment, after which the reduction in absolute risk gradually increased with the duration of follow up.[9]

Care must be taken in calculating absolute risk differences in meta-analysis programs. For example, most programs require the number of participants in each arm of the trial to be input as denominators. Since trials tend to have differing lengths of follow up, pooled absolute differences calculated using participants, rather than person years as denominators, will assume

equal length of follow up across trials, and result in false estimates of absolute risk differences.

Interpretation
In calculating an overall number needed to treat for a meta-analysis, all the data from the trials are taken and pooled—producing a less useful result than that provided by the individual trials. In terms of health economics, an incremental cost effectiveness analysis of an intervention at different levels of baseline risk will almost always be more informative than a summary of cost effectiveness based on a pooled number needed to treat.[22] The pooled value may also result in wrong decisions about who should receive treatment if the concept of a threshold number needed to treat, separating those who are likely to benefit from those who are not, is applied.

Deriving number needed to treat
It is preferable to derive numbers needed to treat by applying the relative risk reductions from trials or meta-analyses to estimates of prognosis from cohort studies (representative of the groups for whom treatment decisions are to made), rather than from the trials and meta-analyses themselves.[23] If the relative risk reduction varies across different baseline risks, prognostic variables and regression techniques can be used to produce an estimate of treatment benefit to patients with different baseline risks.[24] This technique can be used for both individual trials[25] and meta-analyses.[26]

Understanding numbers needed to treat
The increasing use of numbers needed to treat is welcome in some respects, but caution is required. Numbers needed to treat are no better understood than other measures.[27] [28] The method of presenting results of studies influences healthcare decisions. Patients,[29] purchasers,[27] general practitioners,[30] and doctors in teaching hospitals[31] are all more likely to believe an intervention is desirable when effectiveness data are presented as a reduction in relative risk than when data from the same studies are presented as a number needed to treat.

Conclusion

In spite of the reservations we outline here, we still believe that the number needed to treat has a place. In the drug treatment of hypertension, numbers needed to treat have been used appropriately to show the greater effectiveness in preventing cardiovascular events achieved when treating older rather than younger patients[32] and treating patients with moderate rather than mild hypertension.[2] When numbers needed to treat are presented for an intervention, the setting in which it occurred, the time period, the

---

Calculating numbers needed to treat

The number needed to treat is the reciprocal of the absolute risk difference for a bad outcome between treated subjects and the control or placebo group—that is, 1 ÷ (risk of bad outcome in placebo group – risk of bad outcome in treated group). It can be calculated by applying the relative risk reduction obtained from a meta-analysis or a trial to a baseline risk that reflects the risk of the type of patients to be treated. For example, statins achieve a pooled relative risk on treatment of 0.69 for all coronary vascular disease events, but the number needed to treat varies according to the risk group of the patients studied as illustrated below.

In a high risk group
The likely risk of a bad outcome in a high risk group of patients might be as high as 5% a year. This can be estimated from studies of prognosis in relevant patient groups. In this case:
Baseline risk of a bad outcome without treatment is $5\% = 0.05$ per year
Risk of a bad outcome on treatment is $0.05 \times 0.69 = 0.0345$ per year
Risk difference is $0.05 - 0.0345 = 0.0155$
Number needed to treat is $1 \div$ risk difference $= 1/0.0155 = 64$ people treated for 1 year to avoid one bad outcome.

In a low risk group
The likely risk of a bad outcome in a low risk group of patients—for example, in primary care—might be as low as 0.5% a year. In this case:
Baseline risk of a bad outcome without treatment is $0.5\% = 0.005$
Risk of a bad outcome on treatment is $0.005 \times 0.69 = 0.0035$
Risk difference is $0.005 - 0.0035 = 0.0016$
Number needed to treat is $1/0.0016 = 645$ people treated for 1 year to avoid one bad outcome

---

outcome, and the baseline risk of the patients for whom the number needed to treat is thought to be applicable should be described.

1 Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;318:1728-33.
2 Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;310:452-4.
3 Rembold CM. Number-needed-to-treat analysis of the prevention of myocardial infarction and death by antidyslipidemic therapy. *J Fam Pract* 1996;42:577-86.
4 Review: statins prevent stroke, especially in patients with coronary heart disease. *Evidence-Based Med* 1998;3:10. (Abstract of: Crouse JR III, Byington RP, Hoen HM, Furberg CD. Reductase inhibitor monotherapy and stroke prevention. *Arch Intern Med* 1997;157:1305-10.)
5 Egger M, Sterne JA, Davey Smith G. Meta-analysis software. 1998. www.bmj.com./archive/7126. (Cited in: Davey Smith G, Egger M. Meta-analysis. Unresolved issues and future developments. *BMJ* 1998;316:221-5.)
6 Cochrane Database of Systematic Reviews. In: Cochrane Collaboration *Cochrane Library*. Issue 2. Oxford: Update Software, 1998.
7 Downs JR, Clearfield M, Weis S, Whitney E, Shapiro DR, Beere PA, et al. Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels. Results of AFCAPS/TexCAPS. *JAMA* 1998;279:1615-22
8 Shepherd J, Cobbe S, Ford E, Isles C, Lorimer A, MacFarlane P, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolaemia. *N Engl J Med* 1995;333:1301-7
9 4S Scandinavian Simvastatin Survival Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian simvastatin survival study (4S). *Lancet* 1994;344:1383-9.
10 Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG, et al. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. *N Engl J Med* 1996;14:1001-9.
11 The Long-term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *N Engl J Med* 1998;339:1349-57.
12 Antiplatelet Trialists' Collaboration. Collaborative overview of randomized trials of antiplatelet therapy. I. Prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. *BMJ* 1994;308:81-106.
13 Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis* 1985;XXVII:335-71
14 Ebrahim S, Davey Smith G. *Health promotion in older people for the prevention of coronary heart disease and stroke*. London: Health Education Authority, 1996.
15 Mulrow C, Lau J, Cornell J, Brand B. Antihypertensive drug therapy in the elderly. *Cochrane Library*. Issue 2. Oxford: Update Software, 1998.
16 Office for National Statistics. *Mortality statistics: cause, 1995*. London: Stationery Office, 1997. (Series DH2;22.)
17 Cucherat M. *EasyMa 1997a. Meta-analysis software*. Lyons: University of Lyons, 1997.
18 Statins. *Bandolier* January 1998;5:2-4. (www.jr2.ox.ac.uk/bandolier/band47/b47-2.html; accessed October 1998).
19 Altman DG. Confidence intervals for the number needed to treat. *BMJ* 1998;317:1309-12.
20 Fleiss JL. The statistical basis of meta-analysis. *Stat Method Med Res* 1993;2:121-45.
21 Egger M, Davey Smith G, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997;315:1533-7.
22 Morris S, McGuire A, Caro J, Pettitt D. Strategies for the management of hypercholesterolaemia: a systematic review of the cost-effectiveness literature. *J Health Serv Res Policy* 1997;2:231-50.
23 Jackson RT, Sackett DL. Guidelines for managing raised blood pressure. *BMJ* 1996;313:64-5.
24 Davey Smith G, Egger M, Phillips AN. Meta-analysis: beyond the grand mean? *BMJ* 1997;315:1610-4.
25 Rothwell PM. Can overall results of clinical trials be applied to all patients? *Lancet* 1995;345:1616-9.
26 Sharp SJ, Thompson SG, Altman DG. The relation between treatment benefit and underlying risk in meta-analysis. *BMJ* 1996;313:735-8.
27 Fahey T, Griffiths S, Peters TJ. Evidence based purchasing: understanding results of clinical trials and systematic reviews. *BMJ* 1995;311:1056-60.
28 McColl A, Smith H, White P, Field J. General practitioners' perceptions of the route to evidence based medicine: a questionnaire survey. *BMJ* 1998;316:361-5.
29 Hux JE, Naylor CD. Communicating the benefits of chronic preventative therapy: does the format of efficacy data determine patients' acceptance of treatment? *Med Decision Making* 1995;15:152-7.
30 Cranney M, Walley T. Same information, different decisions: the influence of evidence on the management of hypertension in the elderly. *Br J Gen Pract* 1996;46:661-3.
31 Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992;117:916-21.
32 Mulrow CD, Cornell JA, Herrera CR, Kadri A, Farnett L, Aguilar C. Hypertension in the elderly. Implications and generalizability of randomized trials. *JAMA* 1994;272:1932-8.

(Accepted 10 February 1999)

Economics notes

Opportunity cost

Stephen Palmer, James Raftery

The concept of opportunity cost is fundamental to the economist's view of costs. Since resources are scarce relative to needs,[1] the use of resources in one way prevents their use in other ways. The opportunity cost of investing in a healthcare intervention is best measured by the health benefits (life years saved, quality adjusted life years (QALYs) gained) that could have been achieved had the money been spent on the next best alternative intervention or healthcare programme.[2]

Opportunity cost can be assessed directly with cost effectiveness or cost utility studies. When two or more interventions are compared cost utility effectiveness analysis makes the opportunity cost of the alternative uses of resources explicit. Cost effectiveness ratios, that is the £/outcome of different interventions, enable opportunity costs of each intervention to be compared.

Although the concept of opportunity cost is fundamental, incorrect conclusions can result from difficulties in applying the concept. Firstly, the study perspective (societal, patient, etc) is critical since it determines which costs and effects to include in the evaluation.[3] A societal perspective incorporates all the costs and benefits regardless of who incurs or obtains them. More restricted perspectives may mask the fact that costs are simply being shifted to another sector rather than being saved.

Secondly, the choice of comparisons can play a crucial part in cost effectiveness analysis, affecting the measurement of opportunity cost. Ideally an intervention should be compared with all relevant interventions, including doing nothing. Without a "do nothing" baseline, the best of two generally undesirable options may be chosen. Sometimes, however, the do nothing option may be unethical, such as when a new treatment is being compared with one that has been shown to be beneficial. Partly for this reason, many studies compare